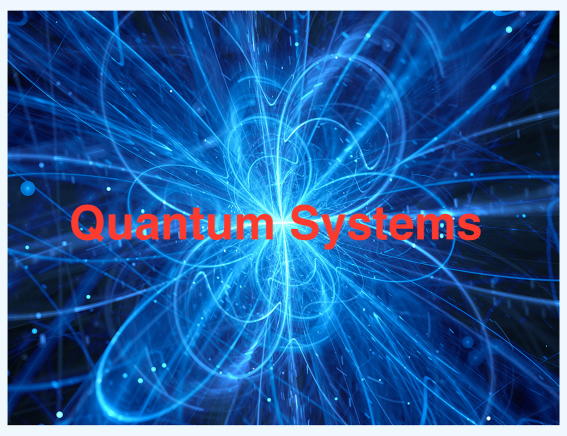
NON-COMMUTATIVE BLAHUT-ARIMOTO ALGORITHMS



NAVNEETH RAMAKRISHNAN, RABAN ITEN,
VOLKHER B. SCHOLZ, AND MARIO BERTA
IMPERIAL COLLEGE LONDON, ETH ZÜRICH AND GHENT UNIVERSITY

SwissMAP
The Mathematics of Physics
National Centre of Competence in Research
FNSNF
SWISS NATIONAL SCIENCE FOUNDATION

MOTIVATION

Many quantities in quantum information such as the channel capacities, mutual information of channels and coherent information of channels are optimization problems over entropic quantities.

How can we efficiently estimate these quantities?

SUMMARY

We generalize alternating optimization algorithms of Blahut-Arimoto type to classical-quantum and fully quantum problems. In particular, we give iterative algorithms to compute

- the **Holevo quantity**: classical capacity of classical-quantum channels,
- the **coherent information** for less noisy channels: quantum capacity of a quantum channel,
- the **quantum mutual information**: the entanglement assisted quantum capacity,
- the **thermodynamic capacity of quantum channels**.

In all cases we provide convergence proofs and analyze the time complexity.

KEY-IDEA:

Alternating optimization algorithms

Goal: Solve (potentially non-convex) optimization problem:

$$\max_{u \in U} f(u)\,.$$

Strategy: We may cleverly recast the problem into a double optimization problem

$$\max_{u \in U} f(u) = \max_{u_1 \in U_1, u_2 \in U_2} g(u_1, u_2)\,.$$

We can then define the iterative procedure by choosing an initial $u_1^{(1)}$ and the following form for the t-th iteration step

$$u_2^{(t)} = \arg\max_{u_2 \in U_2} g(u_1^{(t)}, u_2)\,,$$
$$u_1^{(t+1)} = \arg\max_{u_1 \in U_1} g(u_1, u_2^{(t)})\,.$$

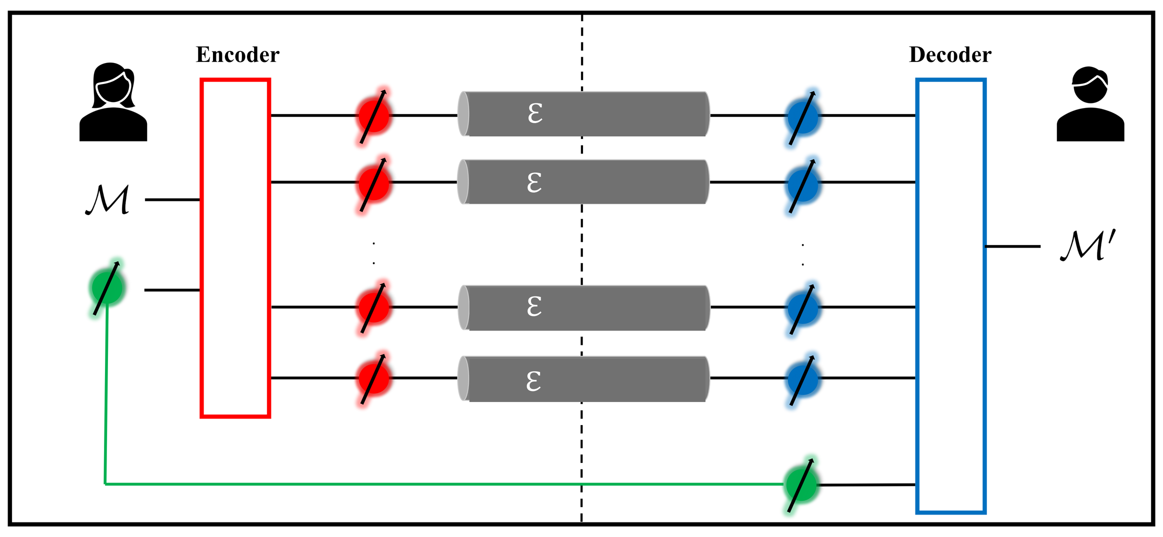
Moreover, we set

$$G^{(t+1)} = g\left(u_1^{(t+1)}, u_2^{(t)}\right)\,.$$

Convergence: Under certain conditions, one can then show

$$\lim_{t \to \infty} G^{(t)} = \max_u f(u)\,.$$

EXAMPLE - QUANTUM MUTUAL INFORMATION



The entanglement assisted classical capacity of a quantum channel $\mathcal{E}$ is given by the mutual information $I(\mathcal{E})$ of the channel, which is defined as

$$I(\mathcal{E}) = \max_\rho S(\rho) + S(\mathcal{E}(\rho)) - S(\mathcal{E}_c(\rho)) = \max_\rho \mathrm{Tr}\,[\rho \mathcal{F}(\rho)]\,, \text{ with}$$
$$\mathcal{F}(\rho) = -\log(\rho) + \mathcal{E}_c^\dagger \log \mathcal{E}_c(\rho) - \mathcal{E}^\dagger \log \mathcal{E}(\rho)\,.$$

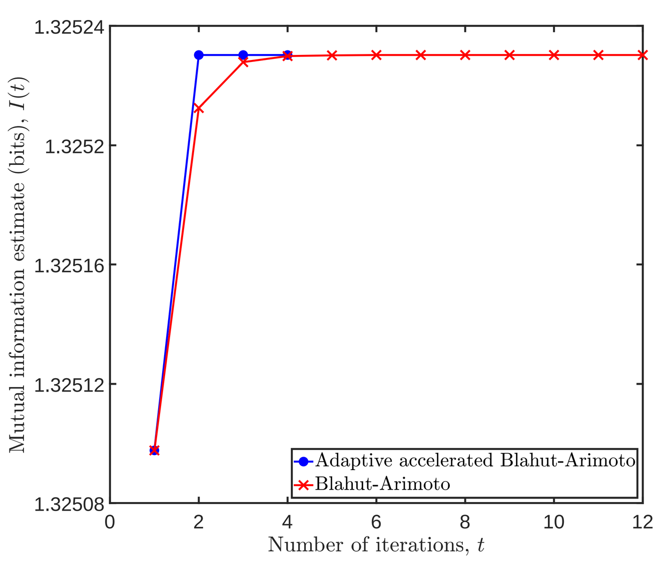
where S denotes the von Neumann entropy and where $\mathcal{E}^c$ denotes the complementary channel of the channel $\mathcal{E}$.

1. We define a two variable extension
$$J(\rho, \sigma) = \mathrm{Tr}\,[\rho \mathcal{F}(\sigma)] - 2D(\rho \| \sigma) \text{ and show that}$$
$$I(\mathcal{E}) = \max_\rho \mathrm{Tr}\,[\rho \mathcal{F}(\rho)] = \max_{\rho,\sigma} J(\rho, \sigma)\,.$$
2. The individual maximizers can be given analytically by
$$\rho^\star(\sigma) = \frac{\exp(\log \sigma + \mathcal{F}(\sigma))}{\mathrm{Tr}\,[\exp(\log \sigma + \mathcal{F}(\sigma))]} \quad \text{and} \quad \sigma^\star(\rho) = \rho\,.$$
3. Start with initial states $\rho^1 = \sigma^1 = I_A/|A|$, where $|A|$ denotes the dimension of the input system of the channel $\mathcal{E}$.
4. We prove that iteratively maximizing $J(\rho, \sigma)$ over ρ and σ converges to $I(\mathcal{E})$ within additive error ϵ after $\frac{2 \log |A|}{\epsilon}$ iterations.

NUMERICS

We consider the amplitude damping channel given by

$$\mathcal{E}_p^{AD}(\rho) = A_0 \rho A_0^\dagger + A_1 \rho A_1^\dagger \text{ with } A_0 = |0\rangle\langle 0| + \sqrt{1-p}|1\rangle\langle 1|, A_1 = \sqrt{p}|0\rangle\langle 1| \text{ for } p \in [0,1].$$



Convergence of the Blahut-Arimoto algorithm to the mutual information of the amplitude damping channel $\mathcal{E}_{0.3}^{AD}$ in the standard and adaptive accelerated case. We evaluate in each iteration step t an *a posteriori* bound until it ensures that the estimation error ϵ satisfies $\epsilon \leq 10^{-6}$ (see our paper on arXiv:1905.01286 for the details about the accelerated case and the *a posteriori* bound).

RESULTS

Channels	Quantity	Time Complexity
$X \to Y$	Mutual information $I(\mathcal{E})$	$\mathcal{O}\left(\frac{\|X\|\|Y\| \log \|X\|}{\varepsilon}\right)$
$X \to B$	Holevo quantity $\chi(\mathcal{E})$	$\mathcal{O}\left(\frac{(\|B\|^3+\|B\|^2\|X\|) \log \|X\|}{\varepsilon} + \|X\|\|B\|^3\right)$
$A \to B$	Thermodynamic capacity $T_\Gamma(\mathcal{E})$	$\mathcal{O}\left(\frac{(\|A\|^3+\|A\|^2\|B\|^2+\|B\|^3) \log \|A\|}{\varepsilon}\right)$
$A \to B$	Coherent information $I_{\mathrm{coh}}(\mathcal{E})$	$\mathcal{O}\left(\frac{(\|A\|^3+\|B\|^3+K^3+\|A\|^2(\|B\|^2+K^2) \log \|A\|}{\varepsilon}\right)$
$A \to B$	Quantum mutual information $I(\mathcal{E})$	$\mathcal{O}\left(\frac{(\|A\|^3+\|B\|^3+K^3+\|A\|^2(\|B\|^2+K^2) \log \|A\|}{\varepsilon}\right)$

Asymptotic worst-case time complexities for an additive ε-approximation. X and Y refer to classical registers while A and B refer to quantum registers. For the coherent information of quantum channels, we require the channel $\mathcal{E}$ to lie in the class of less noisy channels. The Kraus rank of $\mathcal{E}$ is denoted by $K \leq |A||B|$.